

Proposed candidate Standards: Genomic Identifiers

CDEs and modeling strategy

Questions to be answered

- What do outside standard bodies say?
- What do our customers (Domain Workspaces) use?
 - What are use cases for these data elements (where are they likely to be used)?
 - Are there other UML models (in caBIG or in other relevant applications) that use a similar model?
- Can this standard be re-used by different applications in different workspaces
- CDEs AND Model Strategy
- Will this standard be useful in the future?
- What do the caBIG compatibility guidelines say (silver, gold)?



What do outside standard bodies say?

- ▶ Need to write up HL7 usage.
- We understand it is not quite appropriate, but we'd like to write it up.



What do our customers (Domain Workspaces) use?

- What are use cases for these data elements (where are they likely to be used)?
- Are there other UML models (in caBIG or in other relevant applications) that use a similar model?
 - Looking at the following UML models:
 - SEED
 - caCORE
 - Cancer Molecular Pages
 - Function Express Server
 - GO Miner
 - PIR
 - Reactome



Can this standard be re-used by different applications in different workspaces?

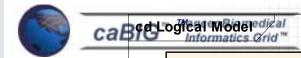
▶ Haven't found any yet. But no reason to believe it will not be.



CDEs AND Model Strategy

- ▶ CDES appear to be very good, well thought out
 - A few questions we had regarding the CDEs
- Modeling strategy may work and maybe something we want to ask developers to do





Gene

- ensemblgeneID: java.lang.String
- genbankAccession: java.lang.String
- genbankAccession.Version: java.lang.String
- entrezgeneID: java.lang.String
- unigeneclusterID: java.lang.String

MessengerRNA

- ensembltranscriptID: java.lang.String
- genbankAccession: java.lang.String
- genbankAccession.Version: java.lang.String
- refseqID: java.lang.String

Protein

- ensemblpeptideID: java.lang.String
- refseqID: java.lang.String
- uniprotkbPrimaryAccession: java.lang.String





